SPARSITY-BASED SOFT DECODING OF COMPRESSED IMAGES IN TRANSFORM DOMAIN

Xianming Liu[1,2], Xiaolin Wu[2], Debin Zhao[1]

[1]School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China, 150001
[2]Department of Electrical & Computer Engineering, McMaster University, Hamilton, Canada, L8S4K1

ABSTRACT

We propose a sparsity-based soft decoding approach to restore compressed images directly in the transform domain of compression (DCT domain specifically examined in this paper). Restoring transform coefficients rather than pixel values prevents the propagation of quantization errors in the image domain. As natural images are statistically non-stationary with spatially varying sparse representations, we develop an adaptive block-wise sparsity-based restoration method that learns and exploits local statistics. Specially, for each DCT block, we collect sample blocks via non-local patch grouping to learn a compact dictionary based on principal component analysis. The resulting block-specific dictionary is used to estimate the corresponding DCT coefficients by a technique of collaborative sparse coding, in which the similarity between sample DCT patches used in dictionary construction is further considered. Experimental results are encouraging and demonstrate that the proposed soft decoding approach performs competitively on restoring compressed images against existing methods.

1. INTRODUCTION

The past decade has witnessed a rapidly growth of research works on sparsity-based image analysis and processing. A number of sparsity-based image restoration techniques were reported [1-3] to deliver superior performances to previous techniques on various application problems such as denoising, super-resolution (upsampling), deconvolution, demosaicking, etc. However, to the best of our knowledge, there is few works that applied the sparsity-based restoration approach to restore compressed images. Ironically, the most common cause of image degradation in practice is compression. Modern image sensors, even the consumer grade digital cameras, offer sufficiently high spatial and spectral resolutions and high signal-to-noise ratio to meet the image quality requirements of most users without any further processing of the raw data. But compression is and will continue to be a vital component of almost all visual communication and computing systems and products, because the sheer volume of image data can easily overwhelm the communication bandwidth and in-device storage. The relative lack of advance in sparsity-based restoration of compressed images is perhaps due to the fact that the compression noises are much more difficult to model than other degradation sources such as blurring and imaging device noises. The non-linearity of quantization operations in image compression systems makes quantization noises signal dependent, far from being white and independent as commonly assumed by works on other image restoration tasks [9].

As motivated above, this work is to investigate if and how well a sparsity-based image restoration approach can be applied to remove or alleviate the adverse effects of quantization in compressed images. In this case, the degraded input image is the decompressed image, which we call hard-decoded image; in contrast, the restored image is called soft-decoded image, and the restoration process is called soft decoding. Soft decoding can be expected to improve the fidelity of a hard-decoded image because all practical image compression methods, including popular international standards JPEG, JPEG 2000, H.264 etc., are not information theoretically optimal. In other words, certain statistical redundancies still exist in the compression code stream. Such residual code redundancies can be exploited, at least theoretically, at the decoder side to improve the fidelity of hard-decoded image by reestimating the original signal based on the knowledge not used by the encoder. For instance, in block-based coding methods like JPEG and H.264, correlations exist between different coding blocks, because natural images have similar local structures due to self-similarity and because the size of the coding blocks cannot be too large due to implementation considerations. These inter-block correlations, which are not exploited by the encoder, can be used by the soft decoder to improve the reconstruction fidelity without spending any extra bits.

In this paper, we focus on soft decoding of DCT domain compressed images for two reasons: 1. common image and video compression standards, such as JPEG and H.264, perform coding in DCT domain; 2. as explained above, the code streams of DCT-based compression methods have certain amount of residual redundancy in the form of inter-block correlation. Unlike most existing image restoration methods, we design soft decoding algorithm to work directly in the trans-

978-1-4799-2341-0/13/$31.00 ©2013 IEEE

form domain instead of pixel domain. This is because inverse DCT transform is required if the restoration is carried out in pixel domain, and it will propagate an isolated quantization error, which is originally confined to a DCT coefficient, to all pixels of the block being restored.

Considering that natural images are statistically non-stationary with spatially varying sparse representations, we perform soft decoding on individual DCT blocks, one at a time, so that the restoration can adapt to local statistics. The locally adaptive sparse representation is constructed out of a dictionary of PCA bases that is learnt using a sample set of approximately matched patches. Even though each DCT coding block is restored individually, the inter-block correlations are accounted for and exploited. This is done by a technique of collaborative sparse coding, in which the similarity between sample DCT patches used in dictionary construction is reused. Specifically, the current DCT block is restored based on the learnt sparse dictionary jointly with all other blocks contributing to the underlying PCA analysis.

The performance of the proposed soft decoding method is further boosted by incorporating the known boundaries of quantizer cells, which is a strong piece of available side information in the DCT code stream, into an objective function of optimal sparse decomposition as constraints.

The rest of this paper is organized as follows. Section 2 introduces the sparsity model and the proposed adaptive dictionary learning method. In Section 3, we detail the proposed sparsity-based soft decoding scheme. Experimental results are given in Section 4. Section 5 concludes the paper.

2. SPARSITY MODEL AND DICTIONARY LEARNING

In block-wise DCT coding, an image is partitioned into nonoverlapping blocks (typically 8×8 pixel blocks). DCT is performed on pixel block independently; the resulting DCT coefficients are scalar quantized according to a quantization table $\mathbf{Q}$. Let $\mathbf{Y}$ be the quantized DCT block, which can be stacked into a vector $\mathbf{y}$ according to the lexicographical order.

2.1. Sparsity Model

Research on image statistics reveals that image patches can be well approximated by a sparse linear combination of elements from an appropriately chosen dictionary. Using this observation as a prior for soft decoding, we seek a sparse representation of each block of DCT coefficients. Let $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_k]$ be the dictionary matrix, where each $\mathbf{d}_i$ represents a basis vector in the dictionary. A DCT patch $\mathbf{y}$ can be represented as a linear combination of atoms in the dictionary $\mathbf{D}$ plus some perturbation ε, that is, $\mathbf{y} = \mathbf{D}\mathbf{a} + \varepsilon, \mathbf{a} \in \Re^{k\times 1}$. We say that the model is sparse if we can achieve $\|\varepsilon\|_2 \ll \|\mathbf{y}\|_2$ and $\|\mathbf{a}\|_0 \ll k$ simultaneously.

2.2. Adaptive Dictionary Learning

Constructing a good dictionary is critical to the performance of the above sparsity model. Since natural images typically exhibit non-stationary statistics, consisting of many heterogeneous regions of significantly different geometric structures or statistical characteristics. Heterogeneous data can be better represented using a mixture of sparsity models, one for each homogeneous subset. Bases for each model are adaptive to the particular homogeneous subset. For this reason, we divide the hard decoded image into a set of overlapped blocks of size 8×8, and perform DCT on these blocks; the resulting DCT coefficient vectors constitute a training data set. In the construction of the sparsity dictionary for restoring DCT coding block $\mathbf{y}_i$, we take advantage of the non-local self-similarity of natural images in learning, and collect similar patches by non-local patch grouping (NLPG) in the training data set. The NLPG procedure guarantees that only the similar sample blocks are used in dictionary learning. The resulting sample set

$$\Psi_i = \{\mathbf{y} \,\big|\, ||\mathbf{y} - \mathbf{y}_i||_2^2/k \le \sigma^2\}, \tag{1}$$

chosen for restoring $\mathbf{y}_i$, where σ is a selection threshold, is subject to principle component analysis. PCA generates the dictionary $\mathbf{D}_i$ whose atoms are the eigenvectors of the covariance matrix of Ψ_i.

3. SOFT DECODING

3.1. Collaborative Sparse Coding

All blockwise DCT-based image/video compression methods suffer from a common problem, that is, sample blocks are encoded independent of each other. Inter-block correlations are totally ignored. This not only reduces the coding efficiency in the first place, but also limits the modeling capability of sparsity-based image prior. The problem is aggravated for low bit rates as vital structural information of the source image is lost or distorted due to the quantization process.

One way of alleviating the above problem is to impose structural sparsity constraints in soft decoding. In the restoration problem sample blocks in Ψ_i are to be estimated simultaneously with $\mathbf{y}_i$, with stipulation that similar blocks are encoded by similar sparsity patterns. Specifically, we explicitly introduce a regularization term into the following optimization problem to preserve the consistency of sparse codes for similar local patches:

$$\min_{\{\alpha_i\}_{i=1}^n} \left\{ \sum_{i=1}^n \|\mathbf{y}_i - \mathbf{D}\mathbf{a}_i\|^2 + \lambda \sum_{i=1}^n \|\mathbf{a}_i\|_1 \right. \\ \left. + \gamma \sum_{i=1}^n \sum_{j=1}^n \|\mathbf{a}_i - \mathbf{a}_j\|_1 W_{ij} \right\}, \tag{2}$$

where n is the number of sample blocks in Ψ_i, W_{ij} measures

Table 1. Performance comparison of tested algorithms in PSNR (dB).

Mehod	QF=10				QF=20				QF=30			
	Butterfly	Lena	Parrots	Bike	Butterfly	Lena	Parrots	Bike	Butterfly	Lena	Parrots	Bike
JPEG	25.32	27.71	29.04	24.06	27.66	29.93	31.61	26.38	29.02	31.19	33.05	27.87
Algorithm [6]	23.77	27.35	28.63	23.77	25.91	29.23	30.91	25.91	27.31	30.45	32.26	27.31
Algorithm [5]	25.71	28.25	29.32	24.28	27.16	29.87	31.01	25.82	27.89	30.67	31.74	26.65
Algorithm [7]	25.21	27.55	28.97	24.01	27.52	29.83	31.54	26.35	28.93	31.17	33.06	27.87
Algorithm [8]	25.66	28.22	29.53	24.25	27.74	30.17	31.91	26.46	29.06	31.36	33.26	27.93
Algorithm [4]	26.03	28.43	29.58	24.43	27.79	30.26	31.94	26.51	29.06	31.38	33.28	27.94
Algorithm [9]	26.25	28.55	29.79	24.60	28.73	30.75	32.36	27.12	30.08	31.90	33.55	28.93
Ours	**26.96**	**28.98**	**30.48**	**25.15**	**29.02**	**30.92**	**32.62**	**27.53**	**30.31**	**32.12**	**33.83**	**29.09**

the similarity between a patches pair $(\mathbf{y}_i, \mathbf{y}_j)$, which is defined as:

$$W_{ij} = \exp\left\{-\frac{||\mathbf{y}_i - \mathbf{y}_j||^2}{\sigma_s^2}\right\}, \quad \sigma_s > 0. \tag{3}$$

The additional regularization term forces that similar blocks have similar sparse representation coefficients.

In addition to the sparsity image prior, the DCT image code stream contains strong pieces of side information on the original image that should be exploited to improve restoration performance. For each DCT coefficient $x(u,v)$, u and v being the indices of the corresponding 2D subband in DCT domain, we know its quantization interval $(q^L_{u,v}, q^U_{u,v})$, i.e.,

$$q^L_{u,v} \le x(u,v) \le q^U_{u,v}. \tag{4}$$

These inequalities can be incorporated into (2) to further confine the solution space and improve the restoration performance. Finally, we formulate our problem of soft decoding as the following constrained convex optimization problem:

$$\begin{aligned} \min_{\{\alpha_i\}_{i=1}^n} & \left\{\sum_{i=1}^n \|\mathbf{y}_i - \mathbf{D}\mathbf{a}_i\|^2 + \lambda \sum_{i=1}^n \|\mathbf{a}_i\|_1 \right. \\ & \left. +\gamma \sum_{i=1}^n \sum_{j=1}^n \|\mathbf{a}_i - \mathbf{a}_j\|_1 W_{ij} \right\}, \\ \text{s.t.,} \quad & \mathbf{q}^L \dot{\le} \mathbf{D}\mathbf{a}_i \dot{\le} \mathbf{q}^U, i = \{1, \cdots, n\} \end{aligned} \tag{5}$$

where $\dot{\le}$ denotes the operation of element-wise comparison, $\mathbf{q}^L$ and $\mathbf{q}^U$ are vectors containing bound values of the quantization interval.

Upon solving (5) and obtaining the optimal sparse coding vectors $\{\mathbf{a}_1^*, \cdots, \mathbf{a}_n^*\}$, we restore the current coding block $\mathbf{y}_i$ to be the weighted average of all reconstructed blocks in the group Ψ:

$$\widehat{\mathbf{y}} = \frac{1}{n} \sum\nolimits_{j=1}^n W_j \mathbf{D}\mathbf{a}_j^*, \tag{6}$$

where W_j is the weight measuring the similarity between the current coding block and other blocks in Ψ, computed in the same way as in Eq.(3).

The above dictionary learning and sparse reconstruction procedure can be carried out iteratively, by using the reconstruction results of the previous iteration as the initial estimates to learn dictionaries for the current iteration.

4. EXPERIMENTAL RESULTS

In this section, experimental results are presented to demonstrate the performance of the proposed soft decoding approach. Our algorithm is compared with existing JPEG deblocking methods: Liu *et al.*'s algorithm [6], Fohoum *et al.*'s algorithm [5], Lee *et al.*'s algorithm [7], Zhai *et al.*'s algorithm [4], and Zhai *et al.*'s algorithm [8]. These methods are included in the comparison group because they can be considered as soft decoding methods for DTC-compressed images. Also included is the well-known denoising algorithm BM3D [9], because the restoration of compressed images can be viewed as a denoising problem, in which the noises are quantization errors.

Table 1 tabulates the PSNR results of the above algorithms on four widely used test images, which are coded by a JPEG coder with quality factors (QF) 10, 20, and 30, respectively. The proposed algorithm has the best PSNR performance for all test images and over all quality factors. Note that the BM3D algorithm needs the knowledge of the variance of noises, and in experiments, we feed the BM3D algorithm the true values of quantization error variances, although in practice this may not always be possible. In this regard, the results of BM3D shown in Table 1 should only be treated as a performance upper bound.

Also, to evaluate the visual qualities of the tested soft decoding algorithms, the Butterfly images restored by the above algorithms are presented in Fig. 1 for readers to judge. It should be clear that the image reconstructed by algorithm [6] is too blurry. The images reproduced by algorithm [5] and algorithm [7] have highly visible noises that accompany edges and textures. Algorithm [8] works better than algorithm

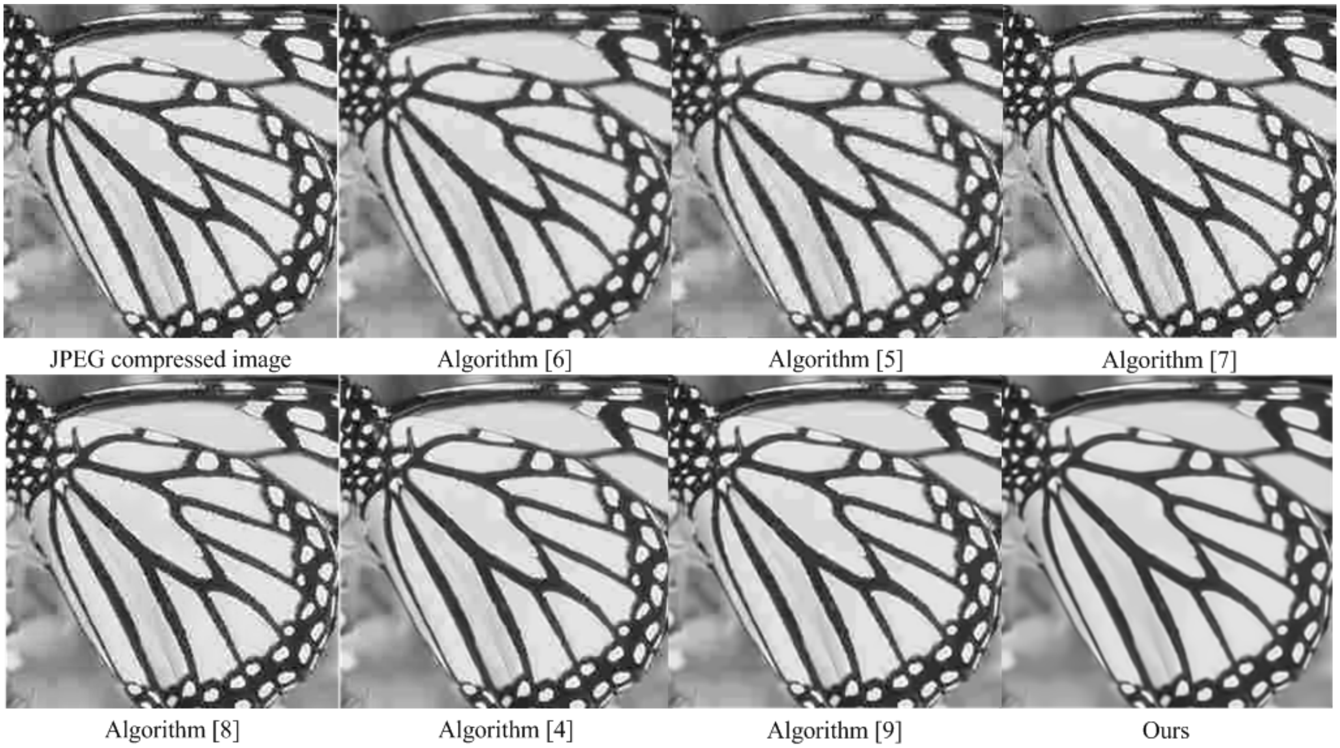

Fig. 1. Comparison of tested methods in visual quality at QF=10.

[5] and [7], however, there are still noticeable artifacts along edges. Algorithm [4] and [9] can suppress most of blocking artifacts and some of the edge-related artifacts, but it still produces some ghosting artifacts along edges, which are particularly visible near corners. The images restored by our soft decoding method are much cleaner, in which the structures and sharpness of edges and textures are well preserved. The proposed method can also remove DCT blocking artifacts in smooth areas completely, and is largely free of the staircase and ringing artifacts along edges.

5. CONCLUSION

We proposed a new sparsity-based soft decoding approach for the restoration of compressed images in DCT domain. The main contribution of this work is the exploitation of inter-block correlations by a technique of collaborative sparse coding. The experimental results are encouraging, opening up the possibility of significantly improving the quality of DCT-compressed images in a postprocess.

6. ACKNOWLEDGEMENT

This work was supported in part by the Major State Basic Research Development Program of China (973 Program) under Grant 2009CB320905 and by the National Science Foundations of China under Grant 61272386.

7. REFERENCES

[1] W. Dong, L. Zhang, G. Shi, X. Wu, "Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization," IEEE Trans. Image Processing 20(7): 1838-1857 (2011)

[2] M. Elad, M. A. T. Figueiredo, Y. Ma, " On the Role of Sparse and Redundant Representations in Image Processing," Proceedings of the IEEE 98(6): 972-982 (2010)

[3] X. Wu, D. Gao, G. Shi, D. Liu, "Color demosaicking with sparse representations," ICIP 2010: 1645-1648

[4] G. Zhai, W. Zhang, X. Yang, W. Lin, Y. Xu, "Efficient Deblocking With Coefficient Regularization, Shape-Adaptive Filtering, and Quantization Constraint," IEEE Trans. on Multimedia 10(5): 735-745 (2008)

[5] A. S. Al-Fohoum and A. M. Reza, "Combined edge crispiness and statistical differencing for deblocking JPEG compressed images," IEEE Trans. Image Process., 10(9):1288-1298(2001).

[6] S. Liu and A. C. Bovik, "Efficient DCT-domain blind measurement and reduction of blocking artifacts," IEEE Trans. Circuits Syst. Video Technol., 12(12): 1139-1149 (2002).

[7] K. Lee, D. S. Kim, and T. Kim, "Regression-based prediction for blocking artifact reduction in jpeg-compressed images," IEEE Trans. Image Processing, 14(1):36-49 (2005).

[8] G. Zhai, W. Zhang, X. Yang, W. Lin, Y. Xu, "Efficient Image Deblocking Based on Postfiltering in Shifted Windows," IEEE Trans. Circuits Syst. Video Techn. 18(1): 122-126 (2008)

[9] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. Image Process., 16(8): 2080-2095 (2007).